

DATABASE

Open Access

TUMIR: an experimentally supported database of microRNA deregulation in various cancers

Lei Dong¹, Min Luo¹, Fang Wang¹, Junwu Zhang¹, Tingting Li² and Jia Yu^{1*}

Abstract

Background: MicroRNAs were found to play an important role in cancers and several literatures exist to describe the relationship between microRNA and cancer, but the expression pattern was still faintly. There is a need for a comprehensive collection and summary of the interactions under experimental support.

Description: TUMIR (<http://www.ncrnalab.com/TUMIR/>), a manually extracted database of experimentally supported microRNA-cancer relationship, aims at providing a large, high-quality, validated comprehensive resource of microRNA deregulation in various cancers. The current version includes a systematic literature search to May-1-2012 using PubMed database, contains data extracted from 205 literatures and 1163 entries describing a regulatory interaction between human microRNAs and cancers. Each entry in the database contains the details of microRNA name, the disease name, case number, control number, p value, the experimentally validated targets, sample type, and a brief description of patients' clinic pathologic parameters mentioned in the same paper. The website has several extensive external links to the related websites and any requests can be made by emailing to tumir_pumc@163.com.

Conclusion: TUMIR is an open access website and will be an accurate clue for the researchers who are interested in better understanding the relationship between miRNAs and cancer.

Keywords: MicroRNA, Cancer, Database

Background

MicroRNAs (miRNAs) are a class of approximately 22–25 nt long endogenous single-stranded RNAs that generally repress the expression of mRNAs. They control mRNA degradation and/or translation through employing a silencing complex to target the 3' UTR region of the mRNAs [1]. In the past decades, miRNAs were found to play an important role in a wide range of biological processes and human diseases especially in tumors [2-4].

Since hundreds of miRNAs were predicted and identified in human cells and tissues, a lot of miRNA-related databases have been built. Of them, several resources provide complete annotation and nomenclature such as miRBase [5], miRGator [6] and miRGen [7]. Several provide experimentally validated microRNA information such as TarBase

[8], miR2Disease [9], dbDEMC [10], miRCancer [11] and miRecords [12]; And also several provide computational predictions such as TargetScan [13], PicTar [14] and microRNA.org [15]. TUMIR is mainly focusing on the data with experimental validation and all samples mentioned in this database must be collected from human. In contrast to TUMIR, the dbDEMC [10] retrieved miRNA-cancer relationship mainly from the analysis of high-throughput expression data. miR2Disease [9], the same as miRCancer [11], integrated data from various experiments without distinguishing human samples from cell lines. Meanwhile, TUMIR is also different from TarBase [8] and miRecords [12] which mainly focused on the miRNA targets. In a word, TUMIR provides a solid evidence of the miRNA deregulation in various cancers.

Accumulating evidence is emerging that miRNAs are closely involved in human cancer initiation, progression, prognosis and diagnosis. Moreover, some miRNAs have been shown to dysregulate and function as tumor suppressors or oncogenes in many kinds of cancers. For example, miR-21 is frequently upregulated in human hepatocellular

* Correspondence: j-yu@ibms.pumc.edu.cn

¹Department of Biochemistry, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences (CAMS) & Peking Union Medical College (PUMC), National Laboratory of Medical Molecular Biology, Beijing 100005, PR China

Full list of author information is available at the end of the article

cancers and lung cancer [16-18], and miR-29b is down-regulated in acute myeloid leukemia [19], and liver cancer [20]. Actually, many groups conducted the same experimental designation to analyze miRNA expression through miRNA array, real-time PCR, northern blot or other methods in human tissues or cells. However, the accuracy of miRNA expression might be interfered by different materials and methods used in these experiments. For example, miR-21, which has been recognized as an oncogene in most of cancers, did not show any differences between case and control groups in some experiments [21]; miR-375, which was found to be down-regulated in multiple cancers, was shown to be up-regulated in some cancers [22-24]. Therefore, finding a clear expression pattern in certain cancer remains a challenging task, especially a systematic documentation and analysis of such experimentally supported miRNA-cancer relationship is needed. To this aim, we have developed a manually extracted database: TUMIR, which provides a large, high-quality, and experimentally validated comprehensive resource of miRNA deregulation in various cancers.

The TUMIR database contains the relationship between miRNAs and cancer, together with the parameters of the design of experiments and patients' clinicopathologies. Additionally, the website has several extensive links, such as a direct link to the paper's information from the NCBI [25] website or a hyper link to the miRNA's information from miRBase [5].

Construction and content

Data sources

We conducted a systematic literature search to May-1-2012 using PubMed database, by means of the terms: "cancer", "carcinoma" and "neoplasm" in combination with the keywords "microRNA" and "miRNA". All results were limited to studies of human. We also searched the reference lists of eligible studies and relevant reviews. Reports fulfilling the following criteria were included if: (i) they were miRNA profiling studies in patients with cancer; (ii) they used cancer and noncancerous samples for comparison; (iii) they did not only use high-throughput methods; (iv) there are more than three samples in the studies and when they got the expression values, the samples could not be pooled. Thus, miRNA profiling studies using only cell lines, or only focusing on the utilization of microarrays, sequencing methods were not included.

Implementation

All the data were stored in a series of relational database implemented with MySQL 5.5.0 on windows system. The front-end web interface implementation is based on PHP and Javascript running under Apache 2.2.21. APIs were generated by java version 1.7.0_15.

The current release of TUMIR contains data extracted from 205 literatures and 1163 entries describing regulatory interactions between human miRNAs and cancer, and complete citations are also provided. The TUMIR data set contains the expression pattern of miRNAs that were up-regulated, down-regulated or no significant changes in various cancers. Each entry in the database contains the details of miRNA name, the disease name, case number, control number, p value, the experimentally validated targets, sample type (tissue, serum, etc.) and a brief description of patients' clinicopathologic parameters mentioned in the same paper. TUMIR provides a user-friendly interface for searching each entry by miRNA name or cancer name.

Utility and discussion

Search page

The TUMIR is designed to serve as a search engine to query of confirmed miRNA-cancer relationship with experimental support. It provides a drop-down list search function. Users can query the database by choosing a miRNA name or cancer type. Once a certain miRNA or cancer type is received as a key word, the search engine will search all the items that contain the query word in TUMIR database. For instance, hsa-miR-375 was reported in a lot of original publications and it was hard to make a full understanding of its expression pattern in various cancer. When users plan to query the expression pattern in all kinds of cancer, they can choose the key word "hsa-miR-375" from the "Search by miRNA Name" drop-down list and retrieve all the systematic documentation of the selected miRNA. After that, a "search results" web page emerges and shows all of information involved in "hsa-miR-375" from queried cancer papers. Each paper is introduced as an individual table, which includes multiple items representing differently experimental data. In the "search results" page, users are able to freely choose the items they are interested in by marking the checked box before the detailed information. They also can click "CHECK_ALL" button located at the top or bottom of the tables to select all of items listed in these tables, or click "CANCEL" button to quit all of selections. However, users have to select at least one item before they press "Submit" button to access the next step. In addition, clicking the "reference" button under the hyperlink of title will lead to the full reference information of the paper (author, date, magazine, etc.). In this page, clicking the "Submit" button will show the summary information of the items you chose. Different expression patterns of miRNAs in cancer are separated by colors. Users will find the relationship between this miRNA and certain cancer by clicking the hyperlinks at the top of tables directly.

Similar to search by miRNA name, users can choose certain cancer from the "Search by Cancer Type" drop-down list, such as "gastric" cancer. Users can also choose their interested items in the "search results" page and click

“Submit” button to find the summary information. Different from searching by miRNA name, the data here is organized by selected cancer and different miRNAs (Figure 1).

Statistics page

The statistics page currently demonstrates 4 histograms of the number of miRNAs and samples associated with multiple cancers archived in TUMIR. These histograms summarized the status of miRNA-cancer researches.

Help page

We employed a user case and a user guide here to make researchers better understand the usability of our website. TUMIR also offers a submission part in this page that allows other researchers to submit un-documented miRNA-cancer patterns. Once accepted by the submission checkup team, the submitted resources will be updated into the database, and will be available to the public in the next release. TUMIR will be updated bimonthly.

Download page

All of the data described above is freely available for download in TUMIR according to the GNU Public License. Users who want to download the current release or previous versions of TUMIR will be taken to the “download” page by clicking “Download” button from the top menu bar. These data are available in multiple formats, such as .sql, .xls and a tab-delimited format. We also

provide a series of APIs here for batch search, researchers can download the APIs, javadocs and a test program here.

Conclusions

The current version of TUMIR attempted to build a widely accessible and user-friendly database that connects miRNAs with cancer through their experimentally supported expression patterns. Although the precise molecular mechanism surrounding many miRNAs and cancer requires further clarification, TUMIR can provide us accurate clue to better describe the relationship between miRNAs and cancer. In order to provide a central resource for the doctors who wish to find some diagnostic markers of certain cancer or for the researchers who wish to study the function and molecular mechanism of miRNAs in some cancer, we not only continuously expand the numbers of known and newly identified items but also create more external links to related websites. We also allow other researchers to contribute to the data contents. Recently, long non-coding RNAs (lncRNAs) in mammals were found to be powerful regulators of gene expression in normal and pathological conditions [26-29]. Similar to miRNAs, lncRNAs have been shown to play vital roles in cancer, and their expression and function are largely unknown [26]. In order to satisfy these requirements, we plan to make rapidly updates and modifications to present a new version of TUMIR.

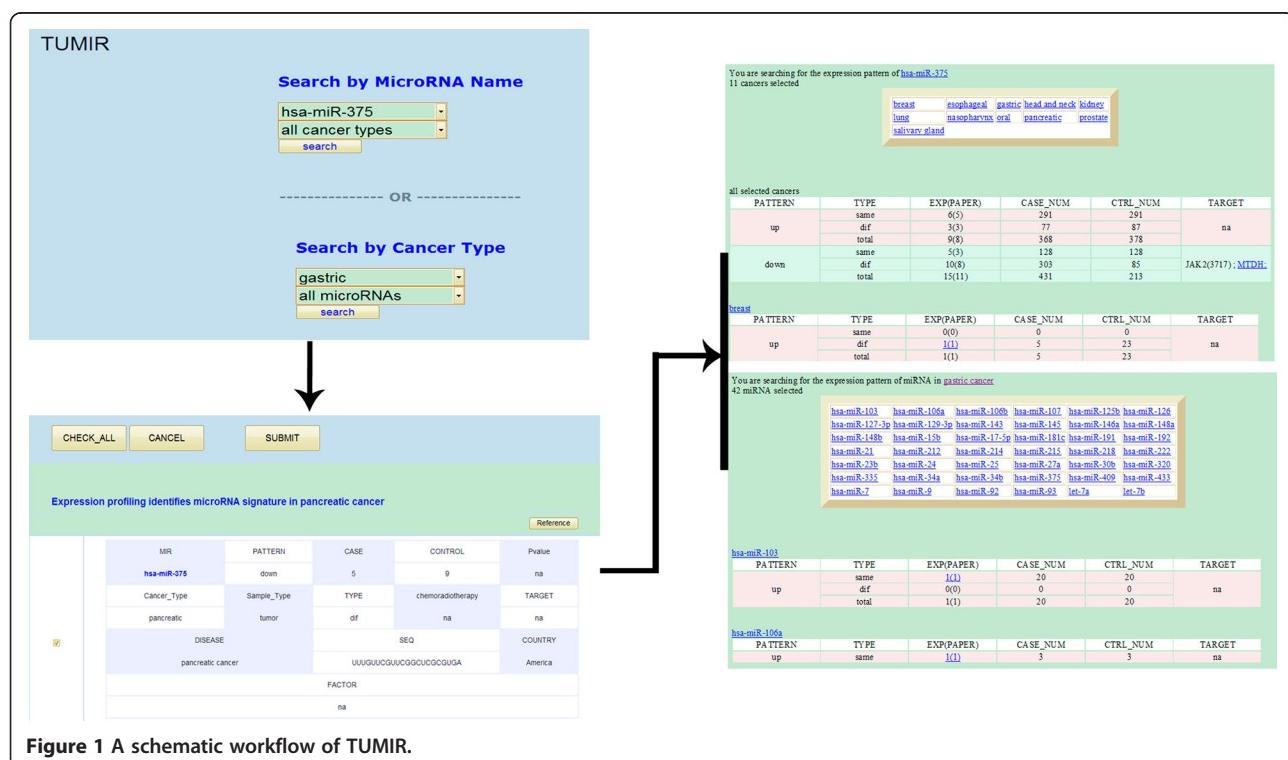


Figure 1 A schematic workflow of TUMIR.

Availability and requirements

Availability

The TUMIR database can be directly accessed from the web page: <http://www.ncrnalab.com/TUMIR/>. This database is publicly and freely accessible, requiring no registration and with no restrictions on use.

Technical requirements

It is recommended that one of the following browsers is used: Mozilla Firefox 3 or Safari 4 on Linux, Mac OSX or Windows, Chrome on Linux or Windows, Internet Explorer 8 on Windows. In order to use the APIs, the java complier compliance level should be 1.7.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

LD, ML, FW and JY performed design and analysis and wrote the manuscript. LD and ML are involved in the data integration and maintenance of the database. T-T L is responsible for bioinformatics analysis. JY, T-T L, J-W Z and LD designed the study. All authors have read and approved the final manuscript.

Acknowledgements

We thank Dr. Changzheng Liu from Peking Union Medical College for valuable suggestion in manuscript writing. This work was supported by grants from the National Natural Science Foundation of China (2011, 91129716, to J.Y.), the National Key Basic Research Program of China (2011CBA01100, to J.Y.), the IBMS, CAMS (2009RC03, to J.Y.; 2010PYB06, to J.Y.), the Beijing Municipal Science & Technology Commission (2010B071, to J.Y.).

Author details

¹Department of Biochemistry, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences (CAMS) & Peking Union Medical College (PUMC), National Laboratory of Medical Molecular Biology, Beijing 100005, PR China. ²Department of Biomedical Informatics, Peking University Health Science Center, Beijing 100191, China.

Received: 22 October 2012 Accepted: 10 April 2013

Published: 17 April 2013

References

1. German MA, Pillay M, Jeong DH, Hetawal A, Luo S, Janardhanan P, Kannan V, Rymarquis LA, Nobuta K, German R, et al: Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nat Biotechnol* 2008, **26**:941–946.
2. Croce CM, Calin GA: MiRNAs, cancer, and stem cell division. *Cell* 2005, **122**:6–7.
3. Calin GA, Croce CM: MicroRNA signatures in human cancers. *Nat Rev Cancer* 2006, **6**:857–866.
4. Zhang W, Dahlberg JE, Tam W: MicroRNAs in tumorigenesis: a primer. *Am J Pathol* 2007, **171**:728–738.
5. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ: MiRBase: tools for microRNA genomics. *Nucleic Acids Res* 2008, **36**:D154–D158.
6. Nam S, Kim B, Shin S, Lee S: MiRGator: an integrated system for functional annotation of microRNAs. *Nucleic Acids Res* 2008, **36**:D159–D164.
7. Megraw M, Sethupathy P, Corda B, Hatzigeorgiou AG: MiRGen: a database for the study of animal microRNA genomic organization and function. *Nucleic Acids Res* 2007, **35**:D149–D155.
8. Sethupathy P, Corda B, Hatzigeorgiou AG: TarBase: a comprehensive database of experimentally supported animal microRNA targets. *RNA* 2006, **12**:192–197.
9. Jiang Q, Wang Y, Hao Y, Juan L, Teng M, Zhang X, Li M, Wang G, Liu Y: miR2Disease: a manually curated database for microRNA deregulation in human disease. *Nucleic Acids Res* 2009, **37**:D98–D104.
10. Zhen Y, Fei R, Changning L, Shummin H, Gang S, Qian G, Lei Y, Yangde Z, Ruoyu M, Ying C, Yi Z, Yang Z, Haitao Z: dbDEMC: a database of differentially expressed miRNAs in human cancers. *BMC Genomics* 2010, **11**(Suppl 4):S5.
11. Boya X, Qin D, Hongjin H, Di W: miRCancer: a microRNA-cancer association database constructed by text mining on literature. *Bioinformatics* 2013, **29**(5):638–644.
12. Xiao F, Zuo Z, Cai G, Kang S, Gao X, Li T: miRecords: an integrated resource for microRNA-target interactions. *Nucleic Acids Res* 2009, **37**:D105–D110.
13. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB: Prediction of mammalian microRNA targets. *Cell* 2003, **115**:787–798.
14. Krek A, Grun D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, et al: Combinatorial microRNA target predictions. *Nat Genet* 2005, **37**:495–500.
15. Betel D, Wilson M, Gabow A, Marks DS, Sander C: The microRNA.org resource: targets and expression. *Nucleic Acids Res* 2008, **36**:D149–D153.
16. Yan LX, Huang XF, Shao Q, Huang MY, Deng L, Wu QL, Zeng YX, Shao JY: MicroRNA miR-21 overexpression in human breast cancer is associated with advanced clinical stage, lymph node metastasis and patient poor prognosis. *RNA* 2008, **14**(11):2348–2360.
17. Lui W, Pourmand N, Patterson BK, Fire A: Patterns of known and novel small RNAs in human cervical cancer. *Cancer Res* 2007, **67**(13):6031–6043.
18. Wu CW, Ng SS, Dong YJ, Ng SC, Leung WW, Lee CW, Wong YN, Chan FK, Yu J, Sung JJ: Detection of miR-92a and miR-21 in stool samples as potential screening biomarkers for colorectal cancer and polyps. *Gut* 2012, **61**(5):739–745.
19. Eyholzer M, Schmid S, Wilkens L, Mueller BU, Pabst T: The tumour-suppressive miR-29a/b1 cluster is regulated by CEBPA and blocked in human AML. *Br J Cancer* 2010, **103**(2):275–284.
20. Xiong Y, Fang JH, Yun JP, Yang J, Zhang Y, Jia WH, Zhuang SM: Effects of microRNA-29 on apoptosis, tumorigenicity, and prognosis of hepatocellular carcinoma. *Hepatology* 2010, **51**(3):836–845.
21. Hites N, Appaiah Chirayu PG, Lida AM, Sunil B, George WS, Yunlong L, Harikrishna N: Persistent upregulation of U6:SNORD44 small RNA ratio in the serum of breast cancer patients. *Breast Cancer Res* 2011, **13**(5):R86.
22. Shen J, Todd NW, Zhang H, Yu L, Lingxiao X, Mei Y, Guarnera M, Liao J, Chou A, Lu CL, Jiang Z, Fang H, Katz RL, Jiang F: Plasma microRNAs as potential biomarkers for non-small-cell lung cancer. *Lab Invest* 2011, **91**(4):579–587.
23. Frezzetti D, De Menna M, Zoppoli P, Guerra C, Ferraro A, Bello AM, De Luca P, Calabrese C, Fusco A, Ceccarelli M, Zollo M, Barbacid M, Di Lauro R, De Vita G: Upregulation of miR-21 by Ras in vivo and its role in tumor growth. *Oncogene* 2011, **30**(3):275–286.
24. Markou A, Tsaroucha EG, Kaklamani L, Fotinou M, Georgoulias V, Lianidou ES: Prognostic value of mature microRNA-21 and microRNA-205 overexpression in non-small cell lung cancer by quantitative real-time RT-PCR. *Clin Chem* 2008, **54**(10):1696–1704.
25. Wheeler DL, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, Church DM, DiCuccio M, Edgar R, Federhen S, et al: Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 2007, **35**:D5–D12.
26. Gutman M, Donaghay J, Carey BW, Garber M, Grenier JK, Munson G, Young G, Lucas AB, Ach R, Bruhn L, Yang X, Amit I, Meissner A, Regev A, Rinn JL, Root DE, Lander ES: lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* 2011, **477**:295–300.
27. Nagano T, Fraser P: No-nonsense functions for long non-coding RNAs. *Cell* 2011, **145**:178–181.
28. Hu W, Yuan B, Flygare J, Lodish HF: Long noncoding RNA-mediated anti-apoptotic activity in murine erythroid terminal differentiation. *Genes Dev* 2011, **25**:2573–2578.
29. Wapinski O, Chang HY: Long noncoding RNAs and human disease. *Trends Cell Biol* 2011, **21**:354–361.

doi:10.1186/2043-9113-3-7

Cite this article as: Dong et al.: TUMIR: an experimentally supported database of microRNA deregulation in various cancers. *Journal of Clinical Bioinformatics* 2013 **3**:7.